



## SEQUENCE LISTING

&lt;110&gt; Consejo Superior de Investigaciones Cientificas

&lt;120&gt; New Gene

&lt;130&gt; DGB/DE/PCT131

&lt;140&gt; PCT/GB99/03019

&lt;141&gt; 1999-09-10

&lt;150&gt; SE 9803069.5

&lt;151&gt; 1998-09-10

&lt;150&gt; US 60/100,873

&lt;151&gt; 1998-09-17

&lt;160&gt; 4

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 2610

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

<223> This gene is referred to in the application as  
Human Death Inducer Obliterator Gene 1. It has now  
been named by the Human Gene Nomenclature  
Committee as Death-Associated Transcription  
Factor (DATF-1)

&lt;400&gt; 1

ctcgggtggcc	gtccgcccac	tccgcggcgt	tccggggaaat	ggctgcgaga	ccctagaggc	60
ctgcggagct	tactccacgg	gaacagcctc	tagataatct	gagttgttga	aaatacgaag	120
cctgttactc	gtgaacagtg	gctgacaaca	gtgttggtgt	gagcctggct	gtctgcttgg	180
acccagaggt	ttcgtctgcc	agggtttttg	gttgatttta	ggatttcagg	gaaaagtgtc	240
caagctttca	gtgttgagc	aggtatggac	gacaaaaggcg	acccgagcaa	tgaggaggca	300
cctaaggcca	tcaaaccac	cagcaaagag	ttcaggaaaa	catgggggtt	tcgaaggacc	360
actatcgcca	agcgagagg	cgcaggggac	gcggaggctg	acccactgga	gccgccaccc	420
ccacagcagc	agctgggcct	gtccctgcgg	cgcagtggga	ggcagcccaa	gcgcactgag	480
cgcgtggagc	agttcctgac	cattgcgcgg	cgccgcggca	ggaggagcat	gcctgtctcc	540
ctggaggatt	ctggtgagcc	cacgtcctgc	cccgccacag	acgccgagac	agcctccgag	600
ggcagcgtgg	aaagcgcttc	tgagaccaga	agcggccccc	agtctgcttc	cacagctgtg	660
aaggaacgac	cagcctcttc	tgaaaagggtg	aaaggagggg	atgaccacga	tgacacctcc	720
gatagtgaca	gcgatggcct	gaccttgaaa	gagcttcaga	atcgccctcg	caggaagcgg	780
gaacaggagc	ccactgagag	gcccctgaaa	gggatccaga	gtcgccctgcg	gaagaagcgc	840
cgggaggagg	gtcccgcgca	gactgtgggc	tccgaggcca	gtgacactgt	ggagggcgtc	900
ctgcccagta	agcaggagcc	cgagaacgat	caggggggttg	tgtcccaggc	tgggaaagat	960
gacagagaga	gtaagttgga	gggaaaggcg	gctcaggaca	tcaaagatga	ggagcctgga	1020
gacttggggc	gaccgaagcc	tgaatgtgag	ggttacgacc	ccaacgccct	gtattgcatt	1080
tgccgccagc	ctcacaacaa	cagggtttatg	atttgctgtg	accgctgtga	agaatggttt	1140
catggcgatt	gtgtgggcat	ttctgaggct	cgaggggaggc	ttttggaaag	gaatggggaa	1200
gactatatct	gccc aaactg	caccattctg	caagtgcagg	atgagactca	ttcagaaacg	1260
gcagatcagc	aggaagctaa	atggagacct	ggagatgctg	atggcaccga	ttgtacaagt	1320

ataggaacaa	tagagcagaa	gtctagcgaa	gaccaaggga	taaagggtag	aattgagaaa	1380
gctgcaaadc	caagtggcaa	gaagaaactc	aagatcttcc	agcctgtgat	agaggcgccct	1440
ggtgcctcaa	aatgtattgg	ccccgggtgc	tgtcacgtgg	cgagccccga	ctcgggtgtac	1500
tgcagtaatg	actgtatcct	caaacacgcc	gcagcgacaa	tgaagtttct	aagctcaggt	1560
aaagaacaga	agccaaagcc	taaagaaaag	atgaagatga	agccagagaa	gccagctctt	1620
ccgaaatgcg	gtgctcaggc	aggtattaaa	atctcttctg	tgacaagag	accagctcca	1680
gaaaaaaaag	agaccacagt	gaagaaggca	gtggtggtcc	ctgcgcggag	tgaagcactc	1740
gggaagggaag	cagcttgtga	gagcagcacg	ccgtcgtggg	cgagcgatca	caattacaat	1800
gcagtaaagc	cagaaaagac	tgctgctccc	tcgccgtcac	tgttgtataa	atgtatgtat	1860
cacctagggg	ttggcctcct	ggaccctctc	cgttctttct	ggatagccat	cccctgggcc	1920
tgtccaggac	tgggagttgc	agctttgtgt	taagctgac	acagacaccg	gctgcacccat	1980
cagcgggaag	cagagcccat	gtccaggatg	cctcctgctg	ccctgtgtcc	atccctagtc	2040
tgtcaggact	tcctgtcact	gttttccaaa	gctgtaaacc	tcactggtga	acgttcacct	2100
taatgattga	ttctttaatc	tctgttttca	ctctcaggct	ctggtaagta	tttgtattct	2160
cttcatccca	gtctgattgc	atagccacac	tgcccgccac	gccacatcca	cccctgtctg	2220
cacatgagtt	gttctgacaa	cagcgtctga	tacgcttcag	tttttccaca	ttgtccacgg	2280
ccagcacatg	aaagcatcac	ttctttttta	tggtgtggga	atctttgcaa	gttagtggtg	2340
catctgattt	tcagggtgtac	atattttttt	gactgggcag	ataggggatt	tttttttttt	2400
tccatgtccg	attcacacgc	tacacaccca	catgaacaca	ttcgaacttc	gaaggcacac	2460
actcctgctt	cataggcccc	acggtaaagt	agttcacacc	tagaactctg	tcctgaccgc	2520
aggacgcgtg	ccttggaactt	ggtattctac	atgtgactgg	ctttcttgcc	ctcgtctctt	2580
gaatgttttag	actcttaaga	tcatatcctg				2610

<210> 2

<211> 2867

<212> DNA

<213> Murine

<400> 2

tccgtggtag	ctctggaaat	ggctgcggga	tccccggcggc	cggggagctt	gtttaagagg	60
cagtccccac	tgtccctttg	gttggttcgaa	gctccggaat	cttctcattg	atgaactgtg	120
actgacaaca	gtggggtgag	gcttggccgt	ctgcttgacc	tggccccagg	tctataattt	180
tatgtaggat	ttccagccaa	aggtttccaa	gctttcagtt	ttgggacagg	tatggatgat	240
aaagggcacc	tgagcaatga	ggaagcacc	aaggctatca	aaccaccag	taaggagttc	300
aggaaaacct	ggggttttcg	aagaaccacg	attgccaaac	gtgagggtgc	aggagacacg	360
gaggcggacc	ccagttagca	gcaaccacag	cagcataacc	tctccctgcg	ccgcagtggg	420
cggcaaccaa	aacgtactga	gagggtagaa	gagtttctta	ccacggttcg	gcgccgaggg	480
aaaaagaatg	tgccggtgtc	cctggaggat	tccagtgagc	ccacatcttc	cacagtcact	540
gatgtggaga	cagcttcgga	ggggagcgtt	gaaagcagtt	ctgagatcag	aagtggccct	600
gtatctgact	ccttagggaa	agaacatcct	gcctcttctg	aaaaggcaaa	aggaggtgaa	660
gaggaagaag	acacctctga	cagtgcagct	gatggcctta	cgttgaaagga	acttcagaac	720
cgcttcgga	gaaagcgaga	gcaagaacct	gtggagaggt	ccctgagagg	cagtcagaat	780
cgcctgagga	agaagcgag	agagggaagat	tctgccgaaa	ctgggagtg	ccaaataggc	840
agtgccgagc	aggacagacc	tctctgtaag	caggagcctg	aggctagtca	gggaccagtg	900
tcccagtcag	agacagatga	catagaaaat	cagttggaag	ggaaggcgac	tcagggaat	960
acagaggaaa	accccgaggga	agcgggcaaa	ccaaagcctg	agtgtgaggt	ttacgacccc	1020
aatgcctgt	actgcatctg	ccgccagcct	cacaacaaca	ggtttatgat	ctgctgtgat	1080
cgggtgtgagg	agtggttcca	tggtgactgt	gtgggtat	ctgaggcccg	agggcggtc	1140
ctggaaagga	acggggaaga	ctacatctgc	ccaaattgca	ccattttgca	agtgcaggat	1200
gagacaaacg	gtagcgccac	caatgagcag	gactctgggt	gcagatctgt	gggtgctgat	1260
ggcacagact	gcacaagcat	aggacagta	gagcagaagt	ccggagaaga	ccagggcata	1320
aagggtagga	ttgagaaggc	agcaaacc	agcggcaaga	aaaaactcaa	gatattccag	1380
cctgtcgtag	aggctcctgg	tgctcctaaa	tgcattggcc	ctgggtgttc	cagtgtagca	1440
cagcctgact	ctgtgtattg	cagtaatgac	tgcattctca	aacacgcagc	agctaccatg	1500
agatttctaa	gttcaggtaa	agaacaaaaa	acaaaaccca	aggaaaaggt	caagacgaag	1560
ccagaaaagt	tcagtcttcc	aaaatgcagt	gttcaggtgg	ggattaaaat	ctcttctgtg	1620

```

cacaagagac tagcgtcaga gaaaagggaa aaccagtgat agaaagtgat gctggcttcc 1680
aggagtgaga cttctgggaa ggaggcagcc tgtgagagca gcacaccatc ctgggcaagt 1740
gaccacaact acaatgctgt gaagccagag aagccagaga agcccactgc actctcgccc 1800
accctattga gtaaatgtac gtatcaccca aaggctggct tcccaggccc ctcccatcat 1860
ctgggtggct gcctggggct gtctaggacc agagtcctgg gtgttctggt gctgatagta 1920
gccagcagct cactgccagc cagaagcaga taccaagatg cctctggacc ccagggtgtt 1980
ctgcctagcc tgtggagcct ctctgggtgg ttcttaaaga gctgtgtagg cctcatgttg 2040
gaggcaattt cttatttcag ttttaggcct tggtaagtat ttgaactgct ctcaacaaga 2100
tgggacaaca tcagtgcact gttcagactc agttcagact tgagttcctc acaggacagc 2160
ccagcacaaa ccatggagtt ttcccacgtt acctattttt cttcaacatg gtcaccactg 2220
ctcttctata gaaaagtgat ttttttcatg agttagggca tttttgaaag gacaggtaga 2280
caatttgtgt ctatttcaca tactacacac ctacataaac aggtttgaat tttgaaggtc 2340
atctgctgct cagatcccat ggtgagtaat cgtgctgacc atacacctac agcagtattc 2400
tcatcaaagt gagaattgtg ggtggtagct tccagactcg tagatgcccc agatttcaaa 2460
tcagtgttta tagtttgagt aaacttttga aacccaagat tagatcactg agtttggagc 2520
cattcttgtt tcaactgctgt acaaggtaca cttttcttac tttttgtaca gtggtttcaa 2580
gtgaattttc atctgtgttc tcatgatagt gttttgttcc atctctgtcc accctcagcc 2640
ccagataagg ggggtctcatg gcctaagctg gccttaaatt ttctgtaaag ttgggggtcc 2700
tcttgcctcc acctgtaaaag agtgctgaaa cttacaggca cacacctgat gtatgctgtg 2760
gggtgtgaac tcagagtgtt ggaccagcag tctaccagct gagctgcagt tctagccatg 2820
gtaactgaaa ctgctgatct tcttgtcttt gtttcctttg tgctggg 2867

```

<210> 3

<211> 562

<212> PRT

<213> Homo sapiens

<400> 3

```

Met Asp Asp Lys Gly Asp Pro Ser Asn Glu Glu Ala Pro Lys Ala Ile
  1             5             10             15

```

```

Lys Pro Thr Ser Lys Glu Phe Arg Lys Thr Trp Gly Phe Arg Arg Thr
          20             25             30

```

```

Thr Ile Ala Lys Arg Glu Gly Ala Gly Asp Ala Glu Ala Asp Pro Leu
          35             40             45

```

```

Glu Pro Pro Pro Pro Gln Gln Gln Leu Gly Leu Ser Leu Arg Arg Ser
          50             55             60

```

```

Gly Arg Gln Pro Lys Arg Thr Glu Arg Val Glu Gln Phe Leu Thr Ile
          65             70             75             80

```

```

Ala Arg Arg Arg Gly Arg Arg Ser Met Pro Val Ser Leu Glu Asp Ser
          85             90             95

```

```

Gly Glu Pro Thr Ser Cys Pro Ala Thr Asp Ala Glu Thr Ala Ser Glu
          100            105            110

```

```

Gly Ser Val Glu Ser Ala Ser Glu Thr Arg Ser Gly Pro Gln Ser Ala
          115            120            125

```

```

Ser Thr Ala Val Lys Glu Arg Pro Ala Ser Ser Glu Lys Val Lys Gly
          130            135            140

```

```

Gly Asp Asp His Asp Asp Thr Ser Asp Ser Asp Ser Asp Gly Leu Thr

```

145	150	155	160
Leu Lys Glu Leu Gln Asn Arg Leu Arg Arg Lys Arg Glu Gln Glu Pro			
	165	170	175
Thr Glu Arg Pro Leu Lys Gly Ile Gln Ser Arg Leu Arg Lys Lys Arg			
	180	185	190
Arg Glu Glu Gly Pro Ala Glu Thr Val Gly Ser Glu Ala Ser Asp Thr			
	195	200	205
Val Glu Gly Val Leu Pro Ser Lys Gln Glu Pro Glu Asn Asp Gln Gly			
	210	215	220
Val Val Ser Gln Ala Gly Lys Asp Asp Arg Glu Ser Lys Leu Glu Gly			
	225	230	235
Lys Ala Ala Gln Asp Ile Lys Asp Glu Glu Pro Gly Asp Leu Gly Arg			
	245	250	255
Pro Lys Pro Glu Cys Glu Gly Tyr Asp Pro Asn Ala Leu Tyr Cys Ile			
	260	265	270
Cys Arg Gln Pro His Asn Asn Arg Phe Met Ile Cys Cys Asp Arg Cys			
	275	280	285
Glu Glu Trp Phe His Gly Asp Cys Val Gly Ile Ser Glu Ala Arg Gly			
	290	295	300
Arg Leu Leu Glu Arg Asn Gly Glu Asp Tyr Ile Cys Pro Asn Cys Thr			
	305	310	315
Ile Leu Gln Val Gln Asp Glu Thr His Ser Glu Thr Ala Asp Gln Gln			
	325	330	335
Glu Ala Lys Trp Arg Pro Gly Asp Ala Asp Gly Thr Asp Cys Thr Ser			
	340	345	350
Ile Gly Thr Ile Glu Gln Lys Ser Ser Glu Asp Gln Gly Ile Lys Gly			
	355	360	365
Arg Ile Glu Lys Ala Ala Asn Pro Ser Gly Lys Lys Lys Leu Lys Ile			
	370	375	380
Phe Gln Pro Val Ile Glu Ala Pro Gly Ala Ser Lys Cys Ile Gly Pro			
	385	390	395
Gly Cys Cys His Val Ala Gln Pro Asp Ser Val Tyr Cys Ser Asn Asp			
	405	410	415
Cys Ile Leu Lys His Ala Ala Ala Thr Met Lys Phe Leu Ser Ser Gly			
	420	425	430
Lys Glu Gln Lys Pro Lys Pro Lys Glu Lys Met Lys Met Lys Pro Glu			
	435	440	445
Lys Pro Ser Leu Pro Lys Cys Gly Ala Gln Ala Gly Ile Lys Ile Ser			

450                      455                      460  
 Ser Val His Lys Arg Pro Ala Pro Glu Lys Lys Glu Thr Thr Val Lys  
 465                      470                      475                      480  
 Lys Ala Val Val Val Pro Ala Arg Ser Glu Ala Leu Gly Lys Glu Ala  
                     485                      490                      495  
 Ala Cys Glu Ser Ser Thr Pro Ser Trp Ala Ser Asp His Asn Tyr Asn  
                     500                      505                      510  
 Ala Val Lys Pro Glu Lys Thr Ala Ala Pro Ser Pro Ser Leu Leu Tyr  
                     515                      520                      525  
 Lys Cys Met Tyr His Leu Gly Val Gly Leu Leu Asp Pro Ser Arg Ser  
                     530                      535                      540  
 Phe Trp Ile Ala Ile Pro Trp Ala Cys Pro Gly Leu Gly Val Ala Ala  
 545                      550                      555                      560  
 Leu Cys

<210> 4  
 <211> 614  
 <212> PRT  
 <213> Murine

<400> 4  
 Met Asp Asp Lys Gly His Leu Ser Asn Glu Glu Ala Pro Lys Ala Ile  
   1                    5                    10                    15  
 Lys Pro Thr Ser Lys Glu Phe Arg Lys Thr Trp Gly Phe Arg Arg Thr  
                     20                    25                    30  
 Thr Ile Ala Lys Arg Glu Gly Ala Gly Asp Thr Glu Ala Asp Pro Ser  
                     35                    40                    45  
 Glu Gln Gln Pro Gln Gln His Asn Leu Ser Leu Arg Arg Ser Gly Arg  
                     50                    55                    60  
 Gln Pro Lys Arg Thr Glu Arg Val Glu Glu Phe Leu Thr Thr Val Arg  
                     65                    70                    75                    80  
 Arg Arg Gly Lys Lys Asn Val Pro Val Ser Leu Glu Asp Ser Ser Glu  
                     85                    90                    95  
 Pro Thr Ser Ser Thr Val Thr Asp Val Glu Thr Ala Ser Glu Gly Ser  
                     100                    105                    110  
 Val Glu Ser Ser Ser Glu Ile Arg Ser Gly Pro Val Ser Asp Ser Leu  
                     115                    120                    125  
 Gly Lys Glu His Pro Ala Ser Ser Glu Lys Ala Lys Gly Gly Glu Glu  
                     130                    135                    140

Glu Glu Asp Thr Ser Asp Ser Asp Ser Asp Gly Leu Thr Leu Lys Glu  
 145 150 155 160  
 Leu Gln Asn Arg Leu Arg Arg Lys Arg Glu Gln Glu Pro Val Glu Arg  
 165 170 175  
 Ser Leu Arg Gly Ser Gln Asn Arg Leu Arg Lys Lys Arg Arg Glu Glu  
 180 185 190  
 Asp Ser Ala Glu Thr Gly Ser Val Gln Ile Gly Ser Ala Glu Gln Asp  
 195 200 205  
 Arg Pro Leu Cys Lys Gln Glu Pro Glu Ala Ser Gln Gly Pro Val Ser  
 210 215 220  
 Gln Ser Glu Thr Asp Asp Ile Glu Asn Gln Leu Glu Gly Lys Ala Thr  
 225 230 235 240  
 Gln Gly Asn Thr Glu Glu Asn Pro Arg Glu Ala Gly Lys Pro Lys Pro  
 245 250 255  
 Glu Cys Glu Val Tyr Asp Pro Asn Ala Leu Tyr Cys Ile Cys Arg Gln  
 260 265 270  
 Pro His Asn Asn Arg Phe Met Ile Cys Cys Asp Arg Cys Glu Glu Trp  
 275 280 285  
 Phe His Gly Asp Cys Val Gly Ile Ser Glu Ala Arg Gly Arg Leu Leu  
 290 295 300  
 Glu Arg Asn Gly Glu Asp Tyr Ile Cys Pro Asn Cys Thr Ile Leu Gln  
 305 310 315 320  
 Val Gln Asp Glu Thr Asn Gly Ser Ala Thr Asn Glu Gln Asp Ser Gly  
 325 330 335  
 Cys Arg Ser Val Gly Ala Asp Gly Thr Asp Cys Thr Ser Ile Gly Thr  
 340 345 350  
 Val Glu Gln Lys Ser Gly Glu Asp Gln Gly Ile Lys Gly Arg Ile Glu  
 355 360 365  
 Lys Ala Ala Asn Pro Ser Gly Lys Lys Lys Leu Lys Ile Phe Gln Pro  
 370 375 380  
 Val Val Glu Ala Pro Gly Ala Pro Lys Cys Ile Gly Pro Gly Cys Ser  
 385 390 395 400  
 Ser Val Ala Gln Pro Asp Ser Val Tyr Cys Ser Asn Asp Cys Ile Leu  
 405 410 415  
 Lys His Ala Ala Ala Thr Met Arg Phe Leu Ser Ser Gly Lys Glu Gln  
 420 425 430  
 Lys Thr Lys Pro Lys Glu Lys Val Lys Thr Lys Pro Glu Lys Phe Ser  
 435 440 445

Leu Pro Lys Cys Ser Val Gln Val Gly Ile Lys Ile Ser Ser Val His  
 450 455 460  
 Lys Arg Leu Ala Ser Glu Lys Arg Glu Asn Pro Val Lys Lys Val Met  
 465 470 475 480  
 Leu Ala Ser Arg Ser Glu Thr Ser Gly Lys Glu Ala Ala Cys Glu Ser  
 485 490 495  
 Ser Thr Pro Ser Trp Ala Ser Asp His Asn Tyr Asn Ala Val Lys Pro  
 500 505 510  
 Glu Lys Pro Glu Lys Pro Thr Ala Leu Ser Pro Thr Leu Leu Ser Lys  
 515 520 525  
 Cys Thr Tyr His Pro Lys Ala Gly Phe Pro Gly Pro Ser His His Leu  
 530 535 540  
 Gly Gly Cys Leu Gly Leu Ser Arg Thr Arg Val Leu Gly Val Leu Val  
 545 550 555 560  
 Leu Ile Val Ala Ser Ser Ser Leu Pro Ala Arg Ser Arg Tyr Gln Asp  
 565 570 575  
 Ala Ser Gly Pro Gln Val Phe Leu Pro Ser Leu Trp Ser Leu Ser Gly  
 580 585 590  
 Trp Phe Leu Lys Ser Cys Val Gly Leu Met Leu Glu Ala Ile Ser Tyr  
 595 600 605  
 Phe Ser Phe Arg Pro Trp  
 610

<210> 5  
 <211> 26  
 <212> DNA  
 <213> Artificial

<220>  
 <223> Primer sequence

<400> 5  
 aggtgtacct tgtacagcag tgaaac

26

<210> 6  
 <211> 16  
 <212> PRT  
 <213> Artificial

<220>  
 <223> Additional N terminal cysteine for addition to a peptide  
 corresponding to amino acids 58-72 of murine DIO-1

<400> 6

Cys Ser Leu Arg Arg Ser Gly Arg Gln Pro Lys Arg Thr Glu Arg Val  
1 5 10 15



# Alignment 1



	1	50
KIAA0333		
SEQ_ID_No3	MDDKGDPSNE EAPKAIKPTS KEFRKTWGFR RTTIKREGA GDAEADPLEP	
Consensus	.....	
	51	100
KIAA0333		
SEQ_ID_No3	PPPQQQLGLS LRRSGRQPKR TERVEQFLTI ARRRGRRSMP VSLEDSGEPT	
Consensus	.....	
	101	150
KIAA0333		
SEQ_ID_No3	SCPATDAETA SEGSVESASE TRSGPQSAST AVKERPASSE KVKGGDDHDD	
Consensus	.....	
	151	200
KIAA0333		
SEQ_ID_No3	TSDSDSDGLT LKELQNLRR KREQEPTERP LKGIQSRLRK KRREEGPAET	
Consensus	.....	
	201	250
KIAA0333		ESKLEG KAAQDIKDEE
SEQ_ID_No3	VGSEASDTVE GVLPSKQEPE NDQGVVSQAG KDDRESKLEG KAAQDIKDEE	
Consensus	.....	....ESKLEG KAAQDIKDEE
	251	300
KIAA0333	PGDLGRPKPE CEGYDPNALY CICRQPHNNR FMICCDRCEE WFHGDCVGIS	
SEQ_ID_No3	PGDLGRPKPE CEGYDPNALY CICRQPHNNR FMICCDRCEE WFHGDCVGIS	
Consensus	PGDLGRPKPE CEGYDPNALY CICRQPHNNR FMICCDRCEE WFHGDCVGIS	
	301	350
KIAA0333	EARGRLERN GEDYICPNCT ILQVQDETHS ETADQQEAKW RPDADGTDC	
SEQ_ID_No3	EARGRLERN GEDYICPNCT ILQVQDETHS ETADQQEAKW RPDADGTDC	
Consensus	EARGRLERN GEDYICPNCT ILQVQDETHS ETADQQEAKW RPDADGTDC	
	351	400
KIAA0333	TSIGTIEQKS SEDQGIKGRI EKAANPSGKK KLKIFQPGPG PVPTQLPVLW	
SEQ_ID_No3	TSIGTIEQKS SEDQGIKGRI EKAANPSGKK KLKIFQP---	
Consensus	TSIGTIEQKS SEDQGIKGRI EKAANPSGKK KLKIFQP...	
	401	450
KIAA0333	QVLEIAVSRS ISAFTHLLHCI SCKVIEAPGA SKCIGPGCCH VAQPDSVYCS	
SEQ_ID_No3	----- ---VIEAPGA SKCIGPGCCH VAQPDSVYCS	
Consensus	..... ...VIEAPGA SKCIGPGCCH VAQPDSVYCS	
	451	500
KIAA0333	NDCILKHAAA TMKFLSSGKE QKPKPKEKMK MKPEKPSLPK CGAQAGIKIS	
SEQ_ID_No3	NDCILKHAAA TMKFLSSGKE QKPKPKEKMK MKPEKPSLPK CGAQAGIKIS	
Consensus	NDCILKHAAA TMKFLSSGKE QKPKPKEKMK MKPEKPSLPK CGAQAGIKIS	
	501	550
KIAA0333	SVHKRPAPPEK KETTVKKAVV VPARSEALGK EAACESSTPS WASDHNYNAV	
SEQ_ID_No3	SVHKRPAPPEK KETTVKKAVV VPARSEALGK EAACESSTPS WASDHNYNAV	
Consensus	SVHKRPAPPEK KETTVKKAVV VPARSEALGK EAACESSTPS WASDHNYNAV	
	551	600
KIAA0333	KPEKTAAPSP SLLYKSTKE- DRRSEEKAAA TAASKKTAPP GSTVGKQPAP	
SEQ_ID_No3	KPEKTAAPSP SLLYKCMYHL GVGLLDPSRS FWIAIPWACP GLGVAALC	
Consensus	KPEKTAAPSP SLLYKcmke. drrle#kaaa faaaiktAcP GlgVaaqc..	

# Alignment 1

	601		650
KIAA0333	RNLVPPKSSF ANVAAATPAI KKPPSGFKGT IPKRPWLSAT PSSGASAARQ		
SEQ_ID_No3			
Consensus	.....		
	651		700
KIAA0333	AGPAPAAATA ASKKFPGSAA LVGAVRKPVV PSVPMASPAP GRLGAMSAAP		
SEQ_ID_No3			
Consensus	.....		
	701		750
KIAA0333	SQPNSQIRQN IRRSLKEILW KRVNDSDDLI MTENEVGKIA LHIEKEMFNL		
SEQ_ID_No3			
Consensus	.....		
	751		800
KIAA0333	FQVTDNRYKS KYRSIMFNLK DPKNQGLFHR VLREEISLAK LVRLKPEELV		
SEQ_ID_No3			
Consensus	.....		
	801		850
KIAA0333	SKELSTWKER PARSVMESRT KLNESKKTa PRQEAIPDLE DSPPVSDSEE		
SEQ_ID_No3			
Consensus	.....		
	851		900
KIAA0333	QQESARAVPE KSTAPLLDVF SSMLKDTTSQ HRAHLFDLNC KICTGQVPSA		
SEQ_ID_No3			
Consensus	.....		
	901		950
KIAA0333	EDEPAPKKQK LSASVKKEDL KSKHDSSAPD PAPDSADEVM PEAVPEVASE		
SEQ_ID_No3			
Consensus	.....		
	951		1000
KIAA0333	PGLESASHPN VDRTYFPGPP GDGHPEPSPL EDLSPCPASC GSGVVTTVTV		
SEQ_ID_No3			
Consensus	.....		
	1001		1050
KIAA0333	SGRDPRTAPS SSCTAVASAA SRPDSTHMVE ARQDVPKPVL TSVMVPKSIL		
SEQ_ID_No3			
Consensus	.....		
	1051		1100
KIAA0333	AKPSSSPDPR YLSVPPSPNI STSESRSPE GDTTLFLSRL STIWKGFINM		
SEQ_ID_No3			
Consensus	.....		
	1101		1150
KIAA0333	QSVAKFVTKA YPVSGCFDYL SEDLPDTIHI GGRIAPKTVW DYVGKLGKSSV		
SEQ_ID_No3			
Consensus	.....		
	1151		1200
KIAA0333	SKELCLIRFH PATEEEEVAY ISLYSYFSSR GRFGVVANN RHVKDLYLIP		
SEQ_ID_No3			
Consensus	.....		
	1201	1226	

## Alignment 1

KIAA0333	LSAQDPVPSK	LLPFEGPGKR	RLSGWR
SEQ_ID_No3			
Consensus	.....	.....	.....

# Alignment 2

1 50  
KIAA0333  
SEQ\_ID\_No4 MDDKGHLSNE EAPKAIKPTS KEFRKTWGFR RTTIKREGA GDTEADPSEQ  
Consensus .....

51 100  
KIAA0333  
SEQ\_ID\_No4 QPQQHNLSLR RSGRQPKRTE RVEEFLTTVR RRGKKNVPVS LEDSSEPTSS  
Consensus .....

101 150  
KIAA0333  
SEQ\_ID\_No4 TVTDVETASE GSVESSEIR SGPVSDSLGK EHPASSEKAK GGEEEDTSD  
Consensus .....

151 200  
KIAA0333  
SEQ\_ID\_No4 SDSA DGLTKE LQNLRRKRE QEPVERSLRG SQNRLRKKRR EEDSAETGSV  
Consensus .....

201 250  
KIAA0333  
SEQ\_ID\_No4 QIGSAEQDRP LCKQEPEASQ GPVSQSETDD IENQLEGKAT QGNTENPRE  
Consensus ..... EnqLEGKAa Qdnk#E#Pr#

251 300  
KIAA0333  
SEQ\_ID\_No4 LGRPKPECEG YDPNALYCIC RQPHNNRFMI CDRCEEFWFH GDCVGISEAR  
Consensus aGrPKPECEg YDPNALYCIC RQPHNNRFMI CDRCEEFWFH GDCVGISEAR

301 350  
KIAA0333  
SEQ\_ID\_No4 GRLLERNGED YICPNCTILQ VQDETHSETA DQGEAKWRPG DADGTDCTSI  
Consensus GRLLERNGED YICPNCTILQ VQDETNGSAT NEQDSGCRSV GADGTDCTSI  
GRLLERNGED YICPNCTILQ VQDEIngeaa ##Q#agcRpg dADGTDCTSI

351 400  
KIAA0333  
SEQ\_ID\_No4 GTIEQKSSSED QGIKGRIEKA ANPSGKKKLK IFQPGPGPVP TQLPVLWQVL  
Consensus GTVEQKSGED QGIKGRIEKA ANPSGKKKLK IFQP-----  
GT!EQKSgED QGIKGRIEKA ANPSGKKKLK IFQP.....

401 450  
KIAA0333  
SEQ\_ID\_No4 EIAVRSISA FTLLHCISCK VIEAPGASKC IGPGCCHVAQ PDSVYCSNDC  
Consensus ----- VVEAPGAPKC IGPGCSSVAQ PDSVYCSNDC  
..... V!EAPGApKC IGPGCchVAQ PDSVYCSNDC

451 500  
KIAA0333  
SEQ\_ID\_No4 ILKHAAATMK FLSSGKEQKP KPKEKMKMKP EKPSLPKCGA QAGIKISSVH  
Consensus ILKHAAATMr FLSSGKEQKp KPKEKmkMkP EKfSLPKCga QaGIKISSVH

501 550  
KIAA0333  
SEQ\_ID\_No4 KRPAPEKKET TVKKAVVPA RSEALGKEAA CESSTPSWAS DHNYNAVKPE  
Consensus KRLApEKrEn pVKK.Vmlaa RSEalGKEAA CESSTPSWAS DHNYNAVKPE

551 600  
KIAA0333  
SEQ\_ID\_No4 KTAAP---SP SLLYKSTKED RRSEKAAAT AASKKTAPPG STVGKQP--A  
Consensus KPEKPTALSP TLLSKCTYHP KAGFPGPSHH LGGCLGLSRT RVLGVLVLIV  
KpaaP...SP sLLsKcTked rageegaaah aagclgaprg rtlGkqp..a

## Alignment 2

	601	650
KIAA0333	PRNLVPPKSS FANVAAATPA IKKPPSGFKG TIPKRPWLSA TPSSGASAAR	
SEQ_ID_No4	ASSSLPARSR YQDASGPQVF LPSLWSLSGW FLKSCVGLML EAIYSFSFRP	
Consensus	arnllParSr %a#aaaaqpa ikklpSgfgg fikkrpgLma eaiSgaSaar	
	651	700
KIAA0333	QAGPAPAAAT AASKKFPGSA ALVGAVRKPV VPSVPMASPA PGRLGAMSAA	
SEQ_ID_No4	W	
Consensus	q.....	
	701	750
KIAA0333	PSQPNSQIRQ NIRRSLKEIL WKRVNDSDDL IMTENEVGKI ALHIEKEMFN	
SEQ_ID_No4		
Consensus	.....	
	751	800
KIAA0333	LFQVTDNRYK SKYRSIMFNL KDPKNQGLFH RVLREEISLA KLVRLKPEEL	
SEQ_ID_No4		
Consensus	.....	
	801	850
KIAA0333	VSKELSTWKE RPARSVMESR TKLHNESKKT APRQEAIPDL EDSPPVSDSE	
SEQ_ID_No4		
Consensus	.....	
	851	900
KIAA0333	EQQESARAVP EKSTAPLLDV FSSMLKDTTS QHRAHLFDLN CKICTGQVPS	
SEQ_ID_No4		
Consensus	.....	
	901	950
KIAA0333	AEDEPAPKKQ KLSASVKKED LKSKHDSSAP DPAPDSADDEV MPEAVPEVAS	
SEQ_ID_No4		
Consensus	.....	
	951	1000
KIAA0333	EPGLESASHP NVDRTYFPGP PGDGHPEPSP LEDLSPCPAS CGSGVVTVT	
SEQ_ID_No4		
Consensus	.....	
	1001	1050
KIAA0333	VSGRDPRTAP SSSCTAVASA ASRPDSTHMV EARQDVPKPV LTSVMVPKSI	
SEQ_ID_No4		
Consensus	.....	
	1051	1100
KIAA0333	LAKPSSSPDP RYLSVPPSPN ISTSESRSPP EGD TTLFLSR LSTIWKGFIN	
SEQ_ID_No4		
Consensus	.....	
	1101	1150
KIAA0333	MQSVAKFVTK AYPVSGCFDY LSEDLPDTH IGGRIAPKTV WDYVGKLKSS	
SEQ_ID_No4		
Consensus	.....	
	1151	1200
KIAA0333	VSKELCLIRF HPATEEEVA YISLYSYFSS RGRFGVVANN NRHVKDLYLI	
SEQ_ID_No4		
Consensus	.....	

## Alignment 2

	1201		1227
KIAA0333	PLSAQDPVPS	KLLPFEGPGK	RRLSGWR
SEQ_ID_No4			
Consensus	.....	.....	.....